Irregular word reading as a marker of semantic decline in Alzheimer's disease: implications for premorbid intellectual ability measurement

Anna Marier[1,2], Mahsa Dadar[1], Florence Bouhali[3], Maxime Montembeault[1*] and for the Alzheimer's Disease Neuroimaging Initiative

Abstract

Background Irregular word reading has been used to estimate premorbid intelligence in Alzheimer's disease (AD) dementia. However, reading models highlight the core influence of semantic abilities on irregular word reading, which shows early decline in AD. The primary objective of this study is to ascertain whether irregular word reading serves as an indicator of cognitive and semantic decline in AD, potentially discouraging its use as a marker for premorbid intellectual abilities.

Method Six hundred eighty-one healthy controls (HC), 104 subjective cognitive decline, 290 early and 589 late mild cognitive impairment (EMCI, LMCI) and 348 AD participants from the Alzheimer's Disease Neuroimaging Initiative were included. Irregular word reading was assessed with the American National Adult Reading Test (AmNART). Multiple linear regressions were conducted predicting AmNART score using diagnostic category, general cognitive impairment and semantic tests. A generalized logistic mixed-effects model predicted correct reading using extracted psycholinguistic characteristics of each AmNART words. Deformation-based morphometry was used to assess the relationship between AmNART scores and voxel-wise brain volumes, as well as with the volume of a region of interest placed in the left anterior temporal lobe (ATL), a region implicated in semantic memory.

Results EMCI, LMCI and AD patients made significantly more errors in reading irregular words compared to HC, and AD patients made more errors than all other groups. Across the AD continuum, as well as within each diagnostic group, irregular word reading was significantly correlated to measures of general cognitive impairment / dementia severity. Neuropsychological tests of lexicosemantics were moderately correlated to irregular word reading whilst executive functioning and episodic memory were respectively weakly and not correlated. Age of acquisition,

---

*Correspondence:
Maxime Montembeault
maxime.montembeault@mcgill.ca
Full list of author information is available at the end of the article

Marier *et al. Alzheimer's Research & Therapy*   (2024) 16:96

Page 2 of 19

a primarily semantic variable, had a strong effect on irregular word reading accuracy whilst none of the phonological variables significantly contributed. Neuroimaging analyses pointed to bilateral hippocampal and left ATL volume loss as the main contributors to decreased irregular word reading performances.

Conclusions While the AmNART may be appropriate to measure premorbid intellectual abilities in cognitively unimpaired individuals, our results suggest that it captures current semantic decline in MCI and AD patients and may therefore underestimate premorbid intelligence. On the other hand, irregular word reading tests might be clinically useful to detect semantic impairments in individuals on the AD continuum.

Keywords Alzheimer's dementia, Premorbid intelligence, Verbal intelligence, Irregular word, Exception word, Reading, Semantic, Neuropsychology, Mild cognitive impairment, Deformation-based morphometry

Background

Alzheimer's disease (AD) is the most common cause of dementia [1]. It is characterized by the insidious accumulation of beta-amyloid and tau proteins ensuing damage to neurons and accompanied with progressive cognitive and behavioral changes. The AD continuum is characterized by three phases, (a) preclinical, (b) mild cognitive impairment (MCI) and (c) AD dementia. One condition which has received increasing attention as an indicator of preclinical AD is subjective cognitive decline (SCD), described as the perception (by oneself or a close contact) of worsening of one's mental abilities, despite seemingly unimpaired performance on objective tests [2]. SCD has been associated with increased risks of future objective cognitive decline [3], as well as increased likelihood of biomarker abnormalities consistent with AD pathology [4]. In the intermediate stage between SCD and AD dementia, MCI patients present objective impairment in one or more cognitive domains [5], but their cognitive changes are mild enough that they require minimal aid or assistance, retaining independence of function in their daily life. On the other hand, AD dementia is associated with more significant cognitive impairments in at least two cognitive domains, which in this case interferes with independence and activities of daily living [6]. The typical amnestic AD dementia most prominently affects learning of new information (episodic memory), but deficits can also be observed in language, visuospatial or executive functions, and through behavioral abnormalities or personality changes. Measuring cognitive decline is therefore central in assessing individuals on the AD continuum. To establish cognitive decline, clinicians will often rely on self and relative-reported changes, as well as comparisons to demographically-adjusted norms of cognitive performance in healthy individuals. Another method is to compare current abilities to an estimate of one's baseline abilities before they were affected by the disease, often referred to as premorbid abilities.

Historically and across many countries, one of the ways to estimate premorbid abilities in patients is the administration of word reading tests [7–22]. This method relies on the assumptions that (a) reading abilities reached by a normal adult is related to their general intelligence and (b) once reading becomes a highly practiced and overlearned skill, it can be maintained at a high level despite deteriorations in other areas of intellectual functioning [23]. These assumptions are consistent with the Cattell-Horn theory of intelligence, which divides general intelligence into two distinct but correlated categories, that is crystallized and fluid intelligence, where crystallized intelligence refers to learned abilities and accumulated knowledge, word reading abilities being an example of, and fluid intelligence refers to more innate mental abilities such as reasoning, memory span and processing speed [24, 25]. As with its name, it is understood that crystallized intelligence remains relatively stable across the lifespan [26] whilst fluid intelligence is vulnerable to the effects of normal ageing [27]. Performances on tasks implicating crystallized intelligence such as reading have therefore been used in older adults or adults with acquired cognitive impairment to estimate general baseline abilities.

In 1978, Nelson and O'Connell introduce the first irregular word reading test, the New (later changed to "National") Adult Reading Test (NART) [28]. The logic behind the use of irregular word reading, as opposed to regular word reading, in estimating premorbid intelligence, is that irregular word reading relies on familiarity to specific words with exceptional spelling. For example, "pint" can only be read correctly by a person who knows of the word and recognises it. Its pronunciation indeed cannot be guessed through the application of common rules of grapheme-phoneme correspondence, as that would only result in reading it like "mint". Therefore, the accurate reading of less frequent irregular words would indicate a larger premorbid vocabulary, which would be related to a high premorbid intellectual quotient (IQ). This assumption was verified on many occasions in healthy adults, most recently when the NART was standardized against the Weschler Adult Intelligence Scale IV, both tests correlating with $r = 0.69$ [29]. Additional to this measure of validity, the reliability of NART-like tests was

Marier *et al. Alzheimer's Research & Therapy*   (2024) 16:96

Page 3 of 19

found to be excellent with an estimated Cronbach's α of around 0.93 [30].

While the validity and reliability of the NART as a measure of premorbid intelligence was clearly demonstrated in cognitively unimpaired adults, the situation might be different in neurodegenerative disorders, whose disease-related cognitive impairments could impact irregular word reading performance. In 1996, Taylor and colleagues pointed out that if estimates of premorbid IQ in patients with neurodegenerative disorders are to be considered valid and accurate, they should (a) not differ significantly from those of demographically matched control subjects (i.e., cognitively unimpaired older adults) and (b) not significantly change as disease progresses in severity.

Regarding Taylor's first criteria, conflicting evidence have been reported. While many cross-sectional studies support the use of NART-like tests in estimating premorbid IQ of AD patients [18, 22, 31–47], many cross-sectional studies have observed significant differences on NART-like test scores between demographically matched HC and AD participants, thus giving support to the theory that irregular word reading might be affected in AD dementia and that this widely used test does not give an accurate estimate of premorbid intelligence in this population [48–61]. Even more importantly, regarding the second criteria, several longitudinal studies observed a significant decline in NART-like performance in AD participants over time [62–70], suggesting that these tests are sensitive to dementia-related cognitive impairments.

These conflicting results could be the result of different factors. A first problem with many of the aforementioned studies is that they have been conducted in the 90s and early 2000s, when concepts like SCD and MCI didn't exist. The same can be said for AD dementia criteria which were not as well developed at the time [6]. In those older studies, it is possible that SCD or MCI participants were classified as normal controls or that other types of dementias were diagnosed as AD dementia. Of note, SCD patients are absent from all the aforementioned studies whilst only three included MCI participants (that is, [22, 42, 70]). Previous studies were also conducted using a relatively low sample size, most often with less than 50 AD participants. This brings particular concern towards the studies in support of the accuracy of irregular word reading premorbid IQ estimates in AD dementia, because in some, nonsignificant differences suggested that a larger sample size would reveal statistical and clinical significance in control-AD comparisons [35, 36, 47]. Nonetheless, when focusing on studies with larger samples sizes and/or longitudinal studies (vs. cross-sectional studies) the evidence seems against the use of irregular word reading as a marker of premorbid IQ in AD

dementia. It is also notable that even in studies supporting their use, NART-like tests were often found to only be accurate at certain, earlier stages of the AD continuum, whilst becoming inaccurate in more severe stages. The stage at which inaccuracies appear varies from study to study, ranging from MCI to moderately severe AD.

Alternatively to the theory that irregular word reading is a measure of premorbid intelligence in AD dementia, some studies suggest that its impairment might reflect a semantic decline [49–52, 56, 62, 63, 67, 68, 71], understood as the loss of general/encyclopedic knowledge. This hypothesis is in line with models of reading that consider the core influence of semantic processes on irregular word reading [72–79]). Consistent with this idea, AD performances on reading and writing tasks that rely to a lesser extent on semantic processing (e.g., reading or writing of words with regular grapheme-phoneme correspondence) appear to be qualitatively more similar to, than divergent from, normal performances, in contrast with tasks requiring semantic processing such as exception word reading [56]. This is further supported by a co-occurring and proportionally similar decline in semantic performances (as measured for instance by picture naming performance) and irregular word reading [67]. Thus, it would appear that a core semantic memory deficit may be the underlying mechanism to impaired irregular word reading in AD dementia, in line with a large body of work suggesting that semantic memory impairments are an early and predominant symptom in MCI and AD dementia [80–84]. Consistently with this hypothesis, the left anterior temporal lobe (ATL) region, involved in semantic processing [85], seems to play a critical role in irregular word reading tasks [86–89], and shows atrophy in patients on the AD continuum [90–92]. Nonetheless, the hypothesis of a semantic deficit causing irregular word reading deficits in the AD continuum remains debated and more evidence is needed to draw solid conclusions regarding the underlying cognitive and neural mechanisms of irregular-word reading in these patients.

The aim of the present article is to assess over a large, well-characterized sample representative of the AD continuum, whether irregular word reading performances (a) significantly differ between diagnostic categories across this continuum and (b) are linked to general cognitive impairment / dementia severity. We hypothesize (1) that demographically-matched MCI and AD participants will perform significantly worse than controls on irregular word reading and (2) that irregular word reading will be correlated with general cognitive impairment / dementia severity. If these two hypotheses are supported by our results and that irregular-word reading performance is not maintained at different stages of AD, we will investigate three additional aims, namely whether

Marier *et al. Alzheimer's Research & Therapy*   (2024) 16:96

Page 4 of 19

the performance on irregular word reading is linked (c) to semantic neuropsychological tests; (d) to psycholinguistic variables associated with semantic processes (but not with psycholinguistic variables associated with phonological processes) and (e) to brain volumes in regions associated with semantic processing. These analyses will contribute to clarify the underlying cognitive and neural mechanisms of irregular word reading deficits. We hypothesize that (3) we will observe a stronger correlation between irregular word reading and tests of semantic processes (e.g. picture naming), as opposed to other tests (e.g., executive functions or episodic memory); (4) the accuracy of single items of the irregular words reading test will be associated with the lexicosemantic variables of the words (e.g., number of sense, semantic neighborhood, concreteness or age of acquisition) as opposed to phonological variables (e.g., number of phonemes, syllables or phonological neighborhood) and (5) finally, we should find neural correlates of semantics to be related to irregular word reading performance, namely the left ATL.

Methods

The data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). ADNI began in 2004 as a public–private partnership under the leadership of Dr. Michael W. Weiner. The primary goal of ADNI has been to detect AD dementia at the earliest possible stage (pre-dementia) and identify ways to track the disease progression. To that end, data from magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers as well as clinical and neuropsychological assessments have been collected to test if they can be combined to measure the progression of the various stages of the AD continuum. The initial five-year study (ADNI-1) was extended by two years in 2009 by a Grand Opportunities grant (ADNI-GO), and in 2011 and 2016 by further competitive renewals of the ADNI-1 grant (ADNI-2, and ADNI-3, respectively). For up-to-date information, see www.adni-info.org.

Participants

Participants over all ADNI studies (1, GO, 2 and 3) who had American National Adult Reading Test (AmNART) scores available at their baseline assessment were included in this study. All participants, aged between 54 and 91 years (inclusive), had completed a minimum of six years of education and did not have vascular dementia, depression, sensory disturbances, or other medical conditions that could interfere with the study. A study-partner who had frequent contact with the participant

(an average of 10 h per week or more) also accompanied them to visits and filled out questionnaires.

Participants were divided into five categories: healthy control (HC), subjective cognitive decline (SCD), early mild cognitive impairment (EMCI), late mild cognitive impairment (LMCI) and Alzheimer's disease (AD) dementia.

The HC status was reserved for participants free of memory complaints, verified by a study partner, beyond what one would expect for age, as well as normal memory function documented by scoring above education adjusted cutoffs on the Logical Memory II subscale (LM II) delayed paragraph recall, from the Wechsler Memory Scale—Revised (WMS-R). Scoring (a) ≥ 9 for 16 or more years of education; (b) ≥ 5 for 8–15 years of education; and (c) ≥ 3 for 0–7 years of education. Additionally, Mini-Mental State Examination (MMSE) score between 24 and 30 (inclusive), Clinical Dementia Rating (CDR) $= 0$, and without significant impairment in activities of daily living. There was no criterion regarding memory complaints.

Participants classified as SCD presented the same scores as HC participants on the WMS-R LM II, MMSE, CDR and presented no significant impairment in activities of daily living. Unlike their HC counterpart, SCD participants presented significant subjective memory concern as reported by subject, study partner, or clinician, as well as significant memory concern confirmed by Cognitive Change Index score ≥ 16.

Participants were classified as EMCI if they presented subjective memory concerns as reported by the subject, their study-partner or clinician, had abnormal memory function documented by scoring within the education adjusted ranges on the WMS-R LM II, scoring inclusively (a) 9–11 for 16 or more years of education; (b) 5–9 for 8–15 years of education; and (c) 3–6 for 0–7 years of education, an MMSE score between 24 and 30 (inclusive) and a CDR score $= 0.5$. Their general cognition and functional performance were sufficiently preserved so that a diagnosis of AD could not be made.

Participants were classified as LMCI if they presented subjective memory concerns as reported by the subject, their study-partner or clinician, had abnormal memory function documented by scoring within the education adjusted ranges on the WMS-R LM II, scoring (a) ≤ 8 for 16 or more years of education; (b) ≤ 4 for 8–15 years of education; and (c) ≤ 2 for 0–7 years of education, an MMSE score between 24 and 30 (inclusive) and a CDR score $= 0.5$. Their general cognition and functional performance were sufficiently preserved so that a diagnosis of AD could not be made.

Diagnosis of AD was made in participants with a memory complaint confirmed by a study partner (or reported

Table 1 Demographics, neuropsychological and language data for all groups

	HC (*n* = 681) 304/377	SCD (*n* = 104) 42/62	EMCI (*n* = 290) 161/129	LMCI (*n* = 589) 362/227	AD (*n* = 348) 196/152	*P* value <.001
Sex, M/F						
Age, years	72.8 ± 6.3[c,e]	72.2 ± 5.6[e]	71.3 ± 7.6[a,d,e]	73.7 ± 7.6 [c]	74.9 ± 7.9[a,b,c]	**<.001**
	55–90	60–90	55–89	54–91	55–91	
Education, years	16.5 ± 2.6[c,,d,e]	16.8 ± 2.5[e]	16.0 ± 2.7[a,e]	16.0 ± 2.8[a,e]	15.3 ± 2.9[a,b,c,d]	**<.001**
Global cognition / Severity						
CDR	0.0 ± 0.0[c,d,e]	0.0 ± 0.0[c,d,e]	0.5 ± 0.0[a,b,e]	0.5 ± 0.0[a,b,e]	0.8 ± 0.3[a,b,c,d]	**<.001**
MoCA	26.2 ± 2.4 (458)[c,d,e]	25.7 ± 2.6 (102)[c,d,e]	24.1 ± 3.0 (287)[a,b,d,e]	22.5 ± 3.2 (234)[a,b,c,e]	17.2 ± 4.5 (173)[a,b,c,d]	**<.001**
MMSE	29.1 ± 1.1[c,d,e]	29.0 ± 1.2[c,d,e]	28.4 ± 1.5[a,b,d,e]	27.2 ± 1.8[a,b,c,e]	23.2 ± 2.0[a,b,c,d]	**<.001**
Episodic memory						
RAVLTdeyaled recall (30 min.)	8.2 ± 3.7 (649)[c,d,e]	8.1 ± 3.6 (99)[c,d,e]	6.5 ± 3.8 (259)[a,b,d,e]	4.3 ± 3.2 (392)[a,b,c,e]	2.5 ± 2.2 (96)[a,b,c,d]	**<.001**
Logical memory, delayed recall	13.3 ± 3.3 [c,d,e]	13.2 ± 3.2[c,d,e]	8.9 ± 1.6[a,b,d,e]	3.9 ± 2.6[a,b,c,e]	1.3 ± 1.8[a,b,c,d]	**<.001**
Language						
Object naming (BNT, 30-item)	28.1 ± 2.1 (399)[d,e]	28.4 ± 1.9[d,e]	27.5 ± 2.6 (289)[d,e]	25.7 ± 4.0 (509)[a,b,c,e]	22.35 ± 5.9 (305)[a,b,c,d]	**<.001**
Semantic fluency (animals)	21.2 ± 5.6[c,d,e]	20.0 ± 5.3[d,e]	18.8 ± 5.1[a,d,e]	16.5 ± 5.0[a,b,c,e]	12.3 ± 5.0[a,b,c,d]	**<.001**
Phonemic fluency (f- words)	14.8 ± 4.7 (461)[c,d,e]	14.1 ± 4.3 (103)[e]	13.7 ± 4.8[a,e]	13.7 ± 4.4 (237)[a,e]	10.9 ± 4.5 (179)[a,b,c,d]	**<.001**
Executive functioning						
Trail making Part-A (seconds)	32.9 ± 11.3[d,e]	34.1 ± 36.8[d,e]	36.8 ± 14.7[d,e]	43.6 ± 21.5[a,b,c,e]	63.6 ± 35.2 (344)[a,b,c,d]	**<.001**
Trail making Part-B (seconds)	78.3 ± 37.4 (678)[c,d,e]	88.3 ± 45.7[d,e]	97.9 ± 50.7 (286)[a,d,e]	125.9 ± 71.8 (581)[a,b,c,e]	196.0 ± 87.0 (315)[a,b,c,d]	**<.001**
Visuospatial functioning						
Copy of a clock	4.7 ± 0.6 (680)[d,e]	4.7 ± 0.6[d,e]	4.6 ± 0.7[d,e]	4.2 ± 1.0[a,b,c,e]	3.4 ± 1.4 (347)[a,b,c,d]	**<.001**
CSF biomarkers						
Beta-amyloid (positive/negative)	162/285 (447)	30/63 (93)	115/144 (259)	263/101 (364)	207/28 (235)	**<.001**
P-tau (positive/negative)	129/313 (442)	33/60 (93)	93/166 (259)	222/141 (363)	191/43 (234)	**<.001**
Irregular word reading						
AmNART (error score)	9.1 ± 8.4[c,d,e]	9.8 ± 7.7[d,e]	12.0 ± 8.7[a,e]	12.9 ± 9.3[a,b,e]	16.5 ± 9.6[a,b,c,d]	**<.001**
	0–48	0–35	0–46	0–46	0–48	

Note. Groups means +/- standard deviation results of demographic, cognitive and language characteristics. Numbers in brackets indicate numbers of participants with the score when less than total. Abbreviations: M: male, F: female, CDR: Clinical Dementia Rating Scale, CSF: cerebrospinal fluid, MoCA: Montreal Cognitive Assessment, MMSE: Mini-Mental State Exam, AVLT: Auditory Verbal Learning Test (which includes a 30-minute delayed recall), BNT: Boston Naming Test, P-Tau: Phosphorylated tau, AmNART: American National Adult Reading Test

a: differed significantly from HC, b: differed significantly from SCD, c: differed significantly from EMCI, d: differed significantly from LMCI, e: differed significantly from AD

only by the study-partner), with abnormal memory function documented by scoring within the education adjusted ranges on the WMS-R LM II, scoring (a) ≤ 8 for 16 or more years of education; (b) ≤ 4 for 8–15 years of education; and (c) ≤ 2 for 0–7 years of education, an MMSE score between 20 and 26 (inclusive), with a CDR score = 0.5 or 1, and who met the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association criteria for probable AD.

Therefore, clinical diagnoses were used to classify patients in the current study. Nonetheless, cerebrospinal fluid (CSF) amyloid- and tau-positivity rates in each group are reported in Table 1. ADNI specific cutoffs, described elsewhere, were as follow: amyloid positive ≤ 977 pg/ml [93], phosphorylated tau positive ≥ 24 pg/ml [94].

In addition to ADNI general inclusion and group classification criteria, we applied for this study two additional specific criteria. The first one was to be native English speakers (excluded *N* = 33). The second criterion was consistency between total AmNART scores and single item-level data on this test, when available, in the ADNI database (excluded *N* = 52). Of the original 2097 and after all considerations, 2012 participants remained, of which 681 HC, 104 SCD, 290 EMCI, 589 LMCI and 348 AD. Demographics of this final sample are provided in the result section.

Marier *et al. Alzheimer's Research & Therapy*   (2024) 16:96

Page 6 of 19

Procedure

Cognitive assessments

AmNART

To measure irregular word reading abilities in an American population, the AmNART (sometimes called ANART) was used. This test is an adaptation of the original British NART [11, 12] developed specifically for the American English population to estimate premorbid intelligence through irregular word reading [10]. The version used by ADNI comprises a list of 50 irregular words, with about half of them identical to the NART. These words are irregular words, also known as exception words, meaning that their actual pronunciation differs from what would be predicted based on the application of grapheme-to-phoneme correspondence (e.g., pint, cellist). They are intended to be printed in order of increasing difficulty and are relatively short to avoid the possible adverse effect of stimulus complexity. Given no time limit, the subject is instructed to read aloud down the list of words, errors made in pronouncing each word is then recorded into an "error score". Participants are allowed to self-correct but are not prompted to do so unless it was difficult to hear what was said and it is necessary to determine whether the pronunciation was correct or incorrect. If they hesitate on two different pronunciations, one correct and the other incorrect, they will be asked which one they think is best.

To assess the involvement of psycholinguistic variables on successful reading of irregular words, we extracted characteristics for each of the 50 AmNART irregular words using the English lexicon project (ELP; [95]) as well as the WordNet [96] data sets, prioritizing ELP, but using WordNet when data was not otherwise available. As control variables, we used (a) word length (number of letters), (b) objective lexical frequency, (c) orthographic neighborhood density and (d) summed bigram frequencies by position. The measure of lexical frequency was obtained from ELP and is the log10 of number of times the word appears in the corpus + 1. The measure of orthographic neighborhood density was the orthographic Levenshtein distance to the 20 closest neighbors in the lexicon (OLD20, [97]). To put it simply, it is a measure of similarity and proximity to other words of the lexicon. Specifically, the OLD20 of a given word is computed as the mean of string edit distances from this word to its 20 closest orthographic neighbors in the lexicon. The edit distance used, Levenshtein distance (LD), corresponds to the number of operations (letter deletion, insertion, or substitution) needed to change a word into another word: for example, the LD from *smile* to *similes* is 2 (two insertions: I and S). Next is summed bigram frequencies by position,

where bigram is defined as a sequence of two letters, it was obtained from ELP and is a measure of frequency of bigrams that is sensitive to positions within words by taking into account the letter positions where the bigram occurs. For example, the bigram frequency for DO in DOG counts DO bigrams only when they appear in the first two positions of a word in the corpus. As lexicosemantic variables, we used (a) age of acquisition, (b) concreteness, (c) number of senses and (d) semantic neighborhood density. The measure of age of acquisition was obtained from ELP, originally recorded by Kuperman and colleagues [98] as the estimated age at which a word was learned, which has been shown to have larger effects in tasks involving semantic information (e.g., picture naming and lexical decision) as opposed to tasks where semantic information was less involved (e.g., reading aloud; [99, 100]). The measure of concreteness was obtained from ELP and is described by Brysbaert and colleagues (2014) [101] as evaluating the degree to which the concept denoted by a word refers to a perceptible, relatable, entity. The measure for number of senses was obtained from WordNet and is described by Miller [96] as the number contexts in which the word can be used to express the number of possible meanings it has. The measure for semantic neighborhood density was obtained from ELP and is described by Mirman and Magnuson [102] as the number and/or proximity of neighboring representations, density referring to how tightly packed the words in the neighborhood are [103]. For phonological variables, we used (a) the number of syllables, (b) the number of phonemes (c) and phonological neighborhood. The measure of phonological neighborhood was obtained from ELP and is, similarly to the aforementioned OLD20, a measure of 20 phonological LD (PLD20).

Mini-mental state exam and Montreal cognitive assessment

To measure general cognitive impairment/ dementia severity, we used scores obtained by participants on the Mini Mental State Examination (MMSE; [104]) and the Montreal Cognitive Assessment (MoCA [105]), two test that are routinely used to screen a wide range of cognitive functions and identify patients on the AD continuum, as well as to determine disease severity.

Boston naming test

To measure lexicosemantic abilities, the Boston Naming Test (BNT) was used [106]. It measures the ability to orally label (name) drawing of objects. Participants have 20 seconds to name what the drawing represents after being presented with the image. A semantic cue is

Marier *et al. Alzheimer's Research & Therapy*   (2024) 16:96

Page 7 of 19

given if the participant fails to recognize the picture (e.g., answering bench instead of tree) or if they state that they do not know what the picture represents. The semantic cue is either a short explanation about the item (e.g., for a mask: "it's part of a carnival fantasy") or a superordinate category (e.g., for a beaver: "it's a kind of animal"). The test presents objects in order of frequency, from most to least common and is discontinued after 6 consecutive failures. ADNI only administers odd numbered items on the standard 60 item BNT, this gives us a maximum score of 30.

Trail making part-B

To measure executive functioning, the trail making test was used. More specifically, we used scores obtained on part-B of the test, which depends on visuomotor, perceptual-scanning skills and requires considerable cognitive flexibility in shifting from number to letter sets under time pressure [107]. 25 circles are presented to the participant which contains numbers 1 through 13 and letters A through L, the circles are scrambled across the given medium, the participant must connect the circles while alternating between numbers and letters in ascending order (e.g., A to 1,1 to B; B to 2; 2 to C), they have up to 300 seconds to complete the test, their time to complete it (in seconds) is recorded as their score.

Rey auditory verbal learning test (30-min delay)

To measure episodic memory, we used the Rey Auditory Verbal Learning Test (RAVLT) [108]. Over five learning trials, participants are read a list of 15 words (list A), they are asked to recall them immediately with no regards for order. After the fifth learning trial, the same task is done using an interfering list (B). Immediately and 30 minutes after administration of list B, list A is recalled, this time without first being read. Scores from the 30-min delay test were used as our measure of episodic memory.

Neuroimaging

All participants received T1-weighted (T1w) MRIs (see http://adni.loni.usc.edu/methods/mri-tool/mri-analysis/ for the detailed MRI acquisition protocols). T1w scans for each participant were pre-processed through our standard pipeline including denoising [109], intensity inhomogeneity correction [110] and intensity normalization into range [0–100]. The pre-processed images were then both linearly (9 parameters: 3 translation, 3 rotation, and 3 scaling; [111]) and nonlinearly [112] registered to a population appropriate average template generated based on 150 ADNI participants. The quality of all the image processing steps, including the linear and nonlinear registrations was visually verified by an experienced rater

(MD). Deformation-based morphometry (DBM) was performed to measure the local anatomical differences in the brains of the participants by estimating the Jacobian determinant of the inverse of the estimated nonlinear deformation field as a proxy of atrophy [113]. DBM values reflect the relative volume of the voxel with respect to the template,i.e. a value of 1 indicates similar volume to the same region in the template, values lower than one indicate volumes smaller than the corresponding region in the template, while values higher than one indicate volumes that are larger than the corresponding region in the template. Therefore, lower DBM values can be interpreted as reduction in the structure volume, i.e., regional atrophy. Voxel-wise DBM maps were used to assess the relationship between brain atrophy and AmNART scores at a voxel level. In addition, mean DBM values within a region of interest (ROI) including the left anterior temporal lobe were used to assess the relationship between atrophy in the left anterior temporal lobe and AmNART scores.

Statistical analyses

Behavioral analyses

To describe the sample, Pearson's chi-square test was used to assess sex as well as amyloid- and tau-positivity differences. One-way analysis of variance (ANOVA) and Tukey post-hoc testing were used for all other variables.

To test the hypotheses that AmNART scores are dementia insensitive and semantic-related we modeled a number of multiple linear regression that predicts AmNART total error score based on (1) diagnostic category, extracting an ANOVA table to test for the factor as a whole, (2) tests of severity (MMSE, MoCA) and (3) neuropsychological tests (BNT, Trail making part-B, RAVLT), controlling for sex, age, and education. The MMSE was also used as a control variable for disease severity when assessing relation to neuropsychological tests, it was favored as results on the MoCA were not available for the whole sample. Epsilon square was used as a measure for effect size (ε^2; [114]). Stein's formula was used to calculate adjusted R^2 [115].

To assess the involvement of each psycholinguistic variable extracted from the AmNART on irregular word reading, we analyzed single-item accuracy with a generalized logistic mixed-effects model using the lme4 package [116]. This analysis was conducted on a subsample of participants who had single item-level AmNART data available, as opposed to only having total AmNART score available (195 HC, 323 LMCI, 156 AD). Single-item accuracy was predicted by length, lexical frequency, orthographic neighborhood, bigram frequencies by position, age of acquisition, concreteness, number of senses, semantic neighborhood density, number of

Marier *et al. Alzheimer's Research & Therapy*   (2024) 16:96

Page 8 of 19

syllables, number of phonemes and phonological neighborhood as fixed effects, with by-item and by-subject random intercepts as random effects. |z| values beyond 1.96 were deemed as significant [117]. Bigram frequencies by position and number of senses were logarithmically transformed to normalize these variables. 20 words with missing values in age of acquisition, objective lexical frequency, concreteness, and/or phonological neighborhood had to be excluded from this analysis. The remaining 30 words were ache, aisle, algae, asthma, blatant, bouquet, cellist, chord, courteous, debt, deny, depot, epitome, façade, gauge, heir, hiatus, hyperbole, naïve, nausea, papyrus, pint, placebo, scion, sieve, simile, subtle, superfluous, thyme and zealot.

Neuroimaging analyses

Similar linear regression models were used to assess the relationship between AmNART scores and voxel-wise DBM values in the subset of the participants that had MRI information available ($N = 1863$), controlling for age, sex, and level of education. A second set of models were also run with diagnostic category as an additional covariate. Voxel-wise results were corrected for multiple comparisons using False Discovery Rate (FDR) controlling technique, with a significance threshold of 0.05.

Second, we conducted a ROI-based analysis to test the specific hypothesis of a relationship between the volume in the left ATL and irregular word reading on a subsample of participants who had neuroimaging data available ($N = 1863$). To do so, we modeled a multiple linear regression that predicts AmNART total error score based on the DBM in the left ATL, controlling for sex, age, education, with and without including diagnostic category as a covariate in the models, similar to the voxel level analyses. The ATL ROI was selected from a previous study [118, 119].

All statistical analyses were performed using R Statistical Software (version 4.2.1; [120]).

Results

Demographic characteristics of the 2012 participants are shown in Table 1. Groups differed with regards to sex χ^2 (4) = 44.94, $p < 0.001$, age F (4, 2007) = 12.15, $p < 0.001$, $\varepsilon^2 = 0.02$ and education F (4, 2007) = 12.57, $p < 0.001$, $\varepsilon^2 = 0.02$. All following analyses were therefore controlled for sex, age and education. Expectedly, groups differed with regards to CSF amyloid- χ^2 (4) = 242.86, $p < 0.001$ and tau-positivity χ^2 (4) = 214.85, $p < 0.001$. Amyloid-positivity rates were 36% in HC, 32% in SCD, 44% in EMCI, 72% in LMCI and 88% in AD participants with available CSF data. Tau-positivity rates were 29% in HC, 35% in SCD, 36% in EMCI, 61% in LMCI and 82% in AD participants with available CSF data. Neuropsychological

and language evaluations broadly revealed the expected patterns of impairment across the AD continuum. First, measures of severity worsened along the continuum of disease progression stages. Second, episodic memory deficits were predominant, but cognitive decline gradually extended to other cognitive domains.

Irregular word reading across the AD continuum

When controlling for sex, age, education, AmNART total error score significantly differed between diagnoses (F [4, 2004] = 52.20 $p < 0.001$, partial $\varepsilon^2 = 0.09$, Fig. 1). Overall, patient groups with more advanced disease progression on the AD continuum made more errors on irregular word reading. Specifically, as seen in Fig. 1, AD dementia participants showed significantly lower performance compared to all other groups. In addition, HC scores also differed significantly from that of EMCI and LMCI. Means and standard deviations of AmNART total scores as well as significant differences are presented in Table 1 (more detailed T ratios, p values and effect sizes for each contrast are presented in Supplementary Table 1).

Of note, we observed the presence of 18 outlier participants whose AmNART total error score deviated by ± 3.29 standard deviation relative to the average of their respective diagnostic group, more precisely 12 HC, 3 EMCI and 3 LMCI. However, excluding these participants did not impact any of the results of the analyses.

Association between irregular word reading and general cognitive impairment / severity

Whole sample and group-specific partial correlations between AmNART and measures of disease severity/ global cognition (MoCA and MMSE) are presented in Fig. 2. These measures control for sex, age and education. Both measures of severity were significantly correlated with total AmNART scores, in all diagnostic groups as well as across the whole sample, further supporting a strong link between AD disease progression and impaired irregular word reading.

Association between irregular word reading and lexicosemantic, executive functioning and episodic memory performance

Whole-sample and group-specific partial correlations between AmNART and the chosen neuropsychological tests (BNT, Trail making part-B and RAVLT delayed recall) are presented in Fig. 3. These measures control for sex, age, education and severity as measured by the MMSE. Total AmNART irregular word reading scores were significantly and moderately correlated with BNT scores (measuring picture naming or lexicosemantic

Marier *et al. Alzheimer's Research & Therapy*   (2024) 16:96

Page 9 of 19
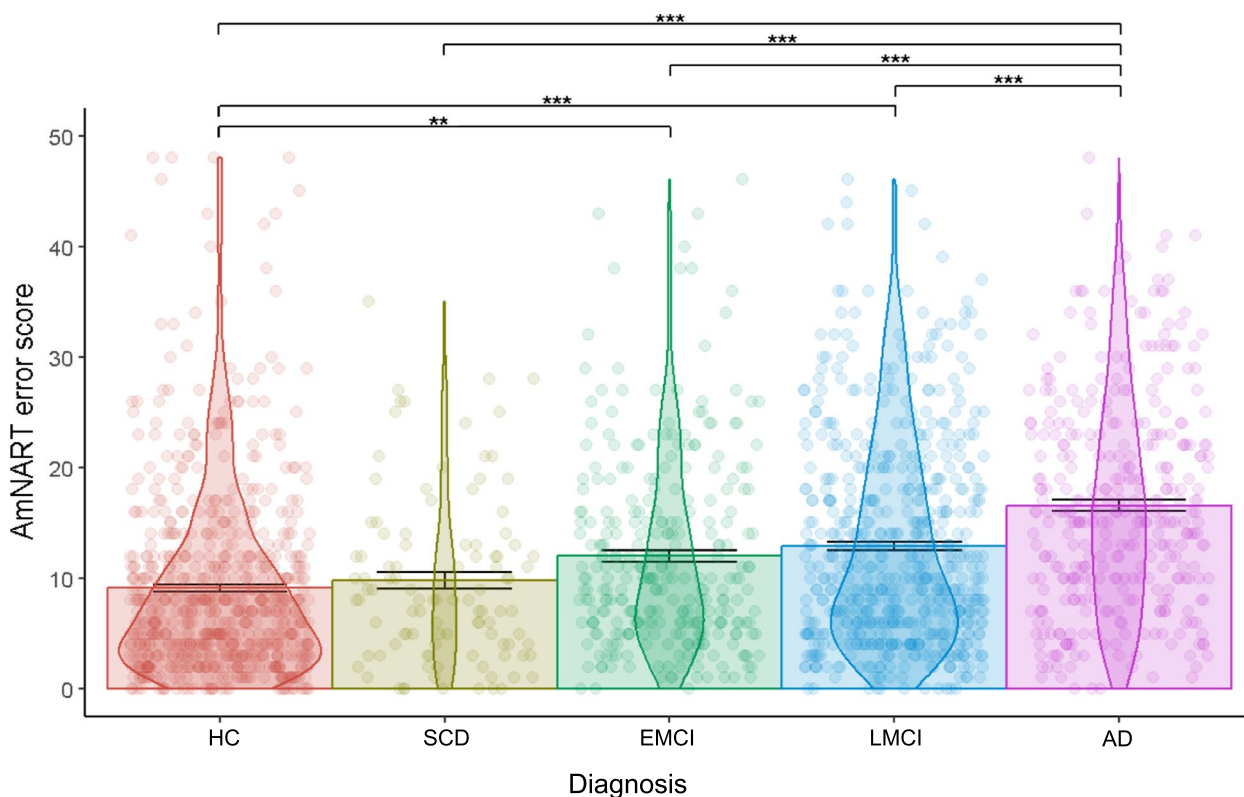
Note. *: p < .05, **: p < .01, ***: p < .001.

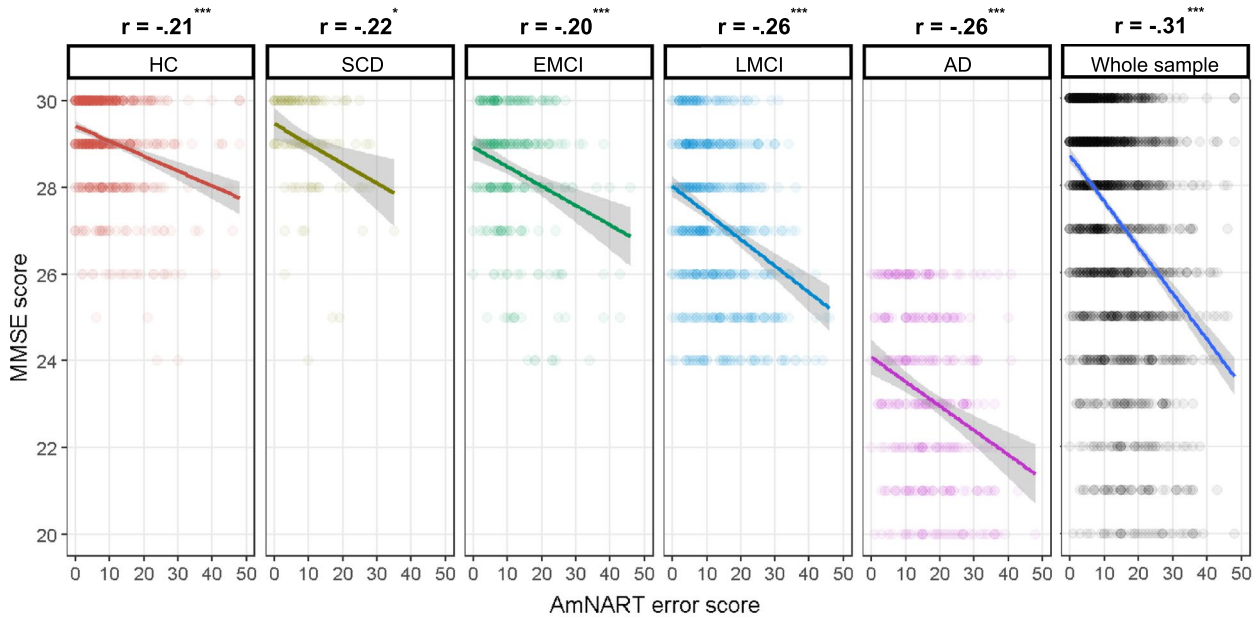**Fig. 1** Relation between AmNART error score and diagnostic category

Table 2 Multiple regression predicting AmNART total error score using neuropsychological tests results

	adjusted R^2	B	SE B	β	P
Model	0.39				<.001
Constant		76.77	4.39		<.001
Control variables					
Sex		-1.15	0.08	-0.13*	<.001
Age		-0.10	0.03	-0.09*	<.001
Education		-2.30	0.43	-0.38*	<.001
Severity: MMSE		-0.64	0.12	-0.15*	<.001
Neuropsychological tests					
Lexicosemantic: Boston Naming Test		-0.79	0.07	-0.31*	<.001
Executive function: Trail making part-B		0.01	0.00	0.06*	<.05
Episodic memory: RAVLTdeyaled recall (30 min.)		0.01	0.06		.887

abilities), weakly but significantly correlated with the Trail making part-B (measuring executive functioning), and poorly correlated with the RAVLT delayed recall (measuring episodic memory), being only significant in the EMCI group ($p < 0.001$) and across the whole sample ($p < 0.05$).

The model created to distinguish between the involvement of lexicosemantic, executive and memory functions in irregular word reading is presented in Table 2. Consistently with the correlational analyses, we observed that the BNT provides a strong contribution to the model (standardized $\beta = -0.31$, $p < 0.001$), the Trail

A. Relation between MMSE and AmNART error score relative to diagnostic category

B. Relation between MoCA and AmNART error score relative to diagnostic category
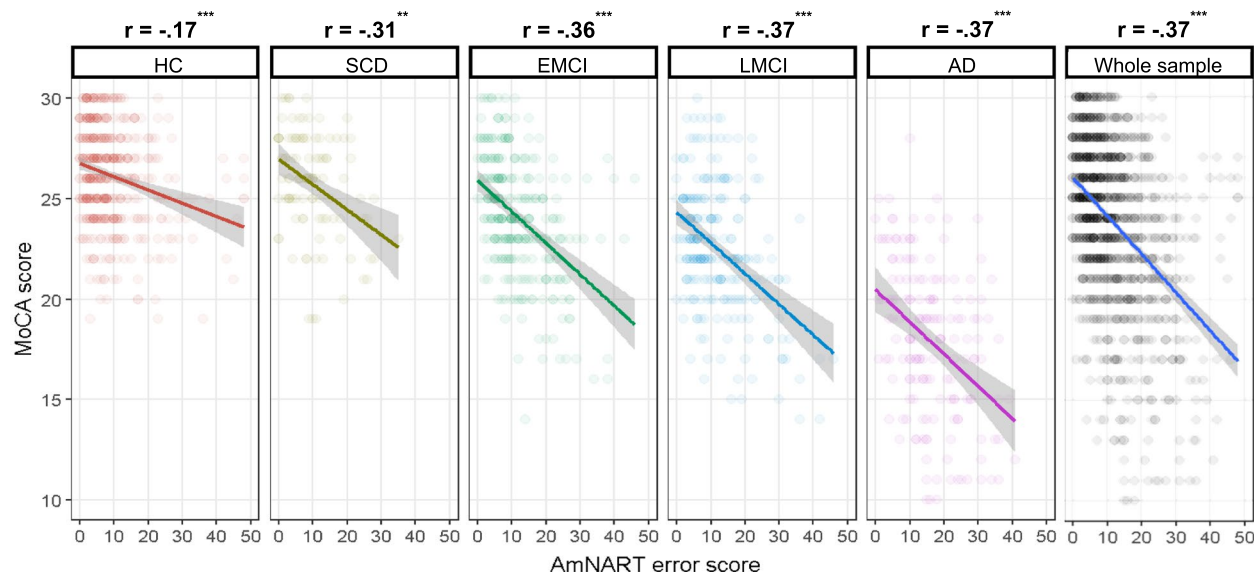
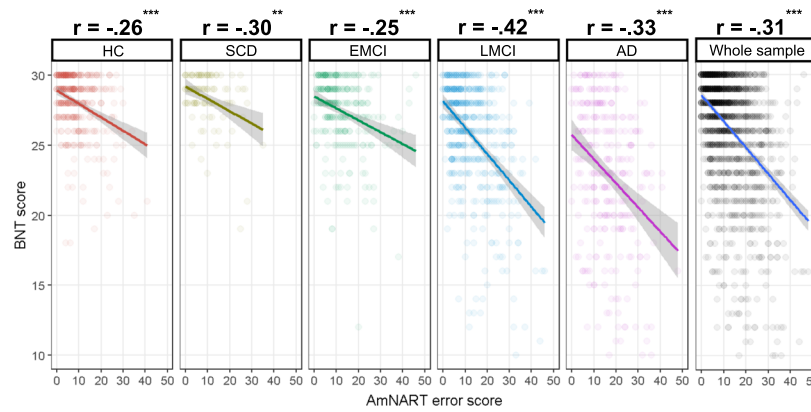*Note.* *: *p* < .05, **: *p* < .01, ***: *p* < .001.

Fig. 2 **A** Relation between MMSE and AmNART error score relative to diagnostic category. **B** Relation between MoCA and AmNART error score relative to diagnostic category

making part-B provides a weak but significant contribution (standardized $\beta = -0.06$, $p < 0.001$) and the RAVLT delayed recall does not provide a significant contribution ($p = 0.887$).

Association between irregular words and psycholinguistic variables (lexicosemantic and phonological)

To better understand the relationships between AmNART irregular words and correct reading, we first selected a subsample for whom single item-level AmNART data was
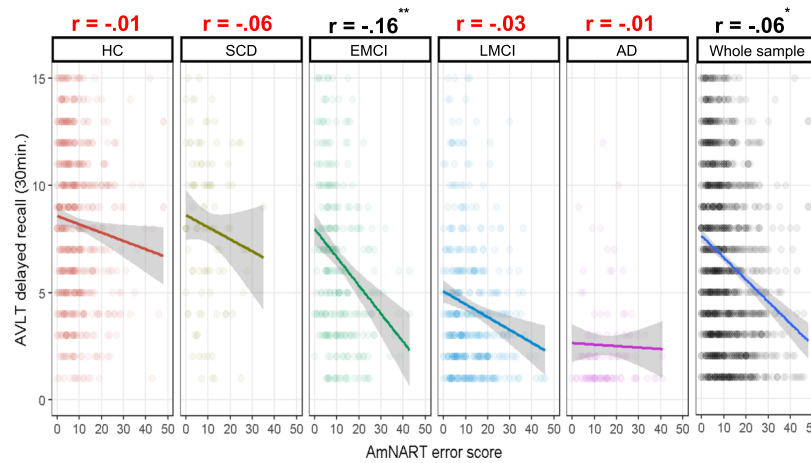
A. Relation between Boston Naming Test and AmNART error score relative to diagnostic category

B. Relation between Trail making part-B and AmNART error score relative to diagnostic category
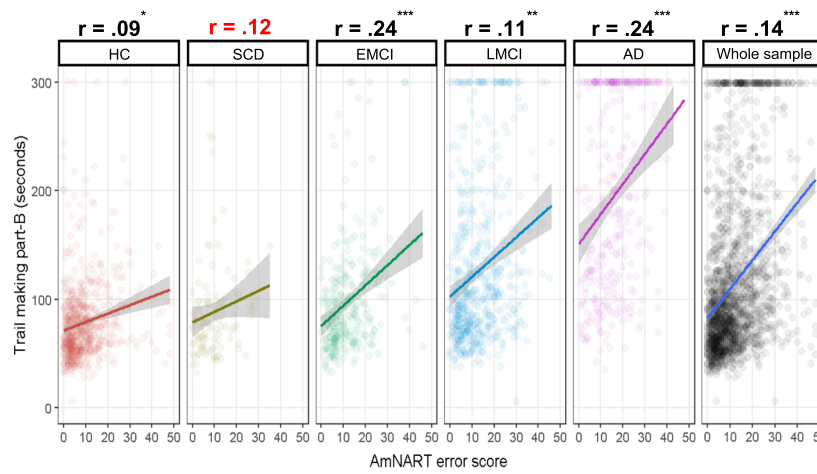
C. Relation between AVLT delayed recall and AmNART error score relative to diagnostic category

Note. *: $p < .05$, **: $p < .01$, ***: $p < .001$.

Fig. 3 **A** Relation between Boston Naming Test and AmNART error score relative to diagnostic category. **B** Relation between Trail making part-B and AmNART error score relative to diagnostic category. **C** Relation between RAVLT delayed recall and AmNART error score relative to diagnostic

Table 3 Generalized logistic mixed-effects model predicting irregular word successful reading using psycholinguistic variables

	B	SE B	z
Control variables			
Length (number of letters)	0.39	0.25	1.54
Frequency (log10)	0.84	0.45	1.89
Orthographic neighborhood (OLD20)	-0.19	0.50	-0.39
Bigram frequencies by position (log)	-0.88	0.37	-2.40
Lexicosemantic variables			
Age of Acquisition	-0.42	0.07	-5.62
Concreteness	-0.07	0.16	-0.46
Number of senses (log)	0.17	0.25	0.69
Semantic neighborhood density	2.40	1.41	1.70
Phonological variables			
Number of syllables	-0.52	0.30	-1.71
Number of phonemes	0.48	0.32	1.49
Phonological neighborhood (PLD20)	-0.55	0.44	-1.24

z values beyond 1.96 in absolute values are deemed as significant [117]

Table 4 Multiple regression predicting AmNART total error score using anterior temporal lobe volume

	adjusted R^2	B	SE B	β	P
Model	0.27				<.001
Constant		36.23	3.00		<.001
Control variables					
Sex		-2.24	0.36	-0.13*	<.001
Age		-0.01	0.02	-0.01*	0.696
Education		-1.41	0.07	-0.44*	<.001
Diagnostic category		1.23	0.13	0.21*	<.001
Region of interest					
Anterior temporal lobe		-3.23	3.00	-0.05*	<.05

available (195 HC, 323 LMCI, 156 AD). The model used to predict irregular word item success based on their psycholinguistic variables is presented in Table 3. While none of the phonological variables had a significant effect on irregular word reading accuracy, there was a significant effect of age of acquisition ($\beta = $ -0.42, $z = $ -5.62).

Link between irregular word reading and brain volumes
Figure 4 shows the results of the significant associations between voxel-wise DBM maps and AmNART scores, including age, sex, and education level as covariates, after correction for multiple comparisons (FDR). At a voxel-wise whole brain level, we observed significant correlations with bilateral medial temporal lobe regions, including the hippocampi, as well as with the ATL, the inferior and middle temporal gyrus, and the fusiform gyrus, predominantly in the left hemisphere. However, no voxels survived FDR correction after including the diagnostic group as covariate. At the ROI level, ATL DBM values were significantly associated with AmNART

scores when including age, sex, and education as covariates (standardized $\beta = $ -0.11, $p < 0.001$). Furthermore, this association remained significant after including diagnostic group as an additional covariate (standardized $\beta = $ -0.05, $p < 0.05$). The model used to predict AmNART error score based on brain volumes in the ATL is presented in Table 4.

Discussion
The present study aimed to assess, over a large and well-characterized sample of participants on the AD continuum, whether irregular word reading performance is an accurate indicator of premorbid intelligence, or a marker of general cognitive and semantic deficits in this population. Results showed that EMCI, LMCI and AD patients make significantly more errors in reading irregular words compared to HC, and that AD patients also make significantly more errors than all other groups. Across the whole AD continuum, as well as within each diagnostic group, irregular word reading abilities were further significantly correlated to measures of general cognitive impairment / dementia severity. This suggests that irregular word reading performances decline throughout the AD continuum, and that even at a finer grain beyond diagnostic categories, a strong link exists between
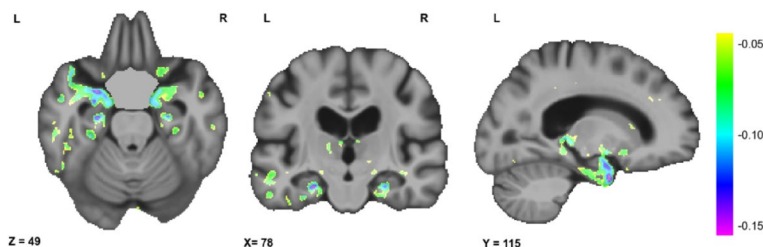

Fig. 4 Relation between voxel-wise DBM maps and AmNART error score. Axial, coronal and sagittal slices showing the t-statistic maps reflecting the significant patterns of brain volume changes in the sample. Colour gradient indicates shrinkage of the tissue (i.e., atrophy). X, Y and Z values indicate MNI coordinates for the displayed slice

dementia severity and irregular word reading difficulties. Furthermore, results indicated significant moderate association between irregular word reading and neuropsychological tests of lexicosemantics, as opposed to weak association to executive function and no association to episodic memory. At the item level, none of the phonological variables had significant effect on irregular word reading accuracy whilst age of acquisition, a semantic variable, provided a significant contribution. Finally, the whole-brain neuroimaging analysis pointed to the hippocampal and left ATL volume loss as the main contributor to decreased irregular word reading performances. These results are consistent with the theory of irregular word reading impairments as an indicator of disease severity and semantic decline, as opposed to an indicator of premorbid IQ in the AD continuum population and pave the way for further investigation on the matter.

Consistent with our first hypothesis, MCI and AD participants performed significantly worse than controls in reading of irregular words, controlling for sex, age and education. EMCI, LMCI and AD participants correctly read an average of 2.9, 3.8 and 7.4 fewer words, respectively, than HC. These measures are comparable to that of Weinborn and colleagues [70] who, when using the Wechsler Test of Adult Reading (WTAR, another 50 irregular word test) found that MCI and AD participants read on average 3.0 and 7.4 fewer words, respectively, than HC. Consistent with hypothesis 2, results indicate that irregular word reading is correlated with general cognitive impairment / dementia severity. This relationship was similar in controls as it was throughout the different diagnostic categories, although and expectedly, that relationship became stronger as we advanced throughout the AD continuum, when larger variations in impairment appeared. Taken together, these two sets of results indicate that irregular word reading performances, although relatively stable throughout normal ageing [121], decline throughout the AD continuum as early as the EMCI stage. Therefore, the assessment of premorbid IQ with the AmNART in participants on the AD continuum violates the criteria set by Taylor and colleagues in 1996 [68], that is to say that an accurate estimate of premorbid IQ should (a) not differ significantly from those of demographically matched control subjects and (b) not significantly change as disease progresses in severity. This has major clinical implications for clinicians and researchers, who could be led to underestimate cognitive changes in people with memory complaints, be more likely to underdiagnose AD continuum conditions or underestimate disease progression in those already diagnosed with one of these conditions. Therefore, it seems preferable for clinicians to rely on comparisons to demographically-adjusted norms of cognitive performance to establish cognitive decline, as well as on repeated measures over time.

Consistent with hypothesis 3, lexicosemantic abilities were the second-best predictor of irregular word reading performances, just after education but largely above dementia severity and other cognitive functions (executive functions and episodic memory). The importance of lexicosemantic abilities in irregular-word reading was further in line with hypothesis 4, as AmNART item success rate was significantly predicted by the age of acquisition of irregular words, which has been associated with semantic representations [99, 100]. This is consistent with the fact that NART-like tests are intended to bypass phonemic decoding by relying more heavily on a person's knowledge of exceptional spelling associated with irregular words. Overall, this set of results highlights the strong association between irregular word reading and semantic abilities, as suggested by Strain and colleagues in 1998 [67] and consistent with the idea of semantic abilities' core influence on irregular word reading performances, particularly but not limited to the AD continuum population. These results are consistent with models of reading that would consider the core influence of semantic abilities on correct reading aloud of irregular words, as emphasized by Taylor and colleagues in their 2015 review [78]. Although not all semantic psycholinguistic variables significantly predicted correct reading, the significant involvement of age of acquisition is consistent with the idea that words acquired at a younger age and used more frequently within the population might be more strongly stored in semantic memory, enhancing the likelihood of successful reading. What these results also show is that executive functioning, episodic memory and phonology do not seem to be as crucial in irregular word reading performance. Overall, these findings suggest that irregular word reading serves as a reliable marker for semantic decline in patients within the AD continuum. Interestingly, earlier studies have suggested that baseline AmNART scores could predict longitudinal cognitive decline in individuals with AD, attributing this effect to the protective nature of premorbid intelligence [66]. However, the current results offer an alternative explanation for this phenomenon: our study aligns more closely with several studies that have demonstrated that baseline semantic memory impairments predict future cognitive decline in AD [122, 123].

Beyond the implication of lexicosemantic abilities as an underlying cognitive mechanism of irregular word reading, we were also interested in looking at the underlying neural mechanisms. Results of our neuroimaging analyses were in line with hypothesis 5. The whole-brain analysis suggested that bilateral hippocampi volumes, as well as with the ATL, the inferior and middle temporal gyrus,

and the fusiform gyrus, predominantly in the left hemisphere, were strongly related to AmNART performance. The significant correlation with the ATL, even when controlling for diagnostic groups, was further confirmed in the ROI-based analysis. The hippocampus is known to be one of the key brain structures affected in AD dementia [124], being the primary site of accumulation of beta-amyloid proteins and phosphorylated tau [125]. Since we found that AmNART scores are also highly correlated with disease severity, it is therefore unsurprising that AmNART scores and hippocampal volumes would be strongly associated. In addition to the well-documented involvement of the hippocampus in episodic memory, research also shows it could well be involved in semantic memory processes ([126–129]). The significant involvement of the ATL in irregular word reading performances in ROI-based analyses is also consistent with previous results supporting the involvement of the ATL in irregular word reading [86–89]. These observations in AD patients are not dissimilar to ATL atrophy in semantic dementia, also accompanied with irregular word reading deficits [79]. However, single word reading tasks like the AmNART have been hypothesized to not be demanding enough on the ATL [77] which could explain the small effect size. Interestingly, the neural correlates of AmNART identified in the current study are notably more specific than the typically reported diffuse pattern of regions (both anterior and posterior) in studies on crystallized intelligence [130, 131]. This provides additional support for the idea that the AmNART may be more sensitive to detecting semantic decline than assessing premorbid intellectual abilities in a population of individuals within the AD continuum.

While the current study fulfills many gaps in the literature (large sample size, well-characterized participants at four different stages on the AD continuum, investigation of underlying cognitive and neural mechanisms of irregular word reading), these results also need to be considered within the context of several limitations. Firstly, the ideal study design would involve obtaining a measure of intelligence before disease onset to correlate with the AmNART score collected during the disease. Unfortunately, this data was unavailable in the ADNI dataset. Secondly, the cross-sectional design of the study does not confirm that irregular word reading declines with time in participants on the AD continuum, as a longitudinal design would. Thirdly, the ADNI cohort is not population-based and underrepresents ethnoculturally diverse populations, its participants are also highly educated and have fewer comorbidities compared to other cohorts [132]. ADNI results must also be interpreted with the caveat that they may have limited external validity and generalizability for more diverse populations. This is a significant problem, especially because racial/ethnic disparities in timeliness and comprehensiveness of dementia diagnosis have already been highlighted [133]. The results of the current study suggest that the AmNART capture semantic decline and might therefore underestimate premorbid intelligence in patients on the AD continuum. The assessment of premorbid function using tools that are not validated in diverse populations might therefore contribute to or even amplify disparities to timeliness of dementia diagnosis. Generalizing to other populations is further complicated by differences in how irregular words are experienced in other languages with more transparent spelling-to-sound correspondences. Italian for example, is more transparent than the more opaque English and is characterized by regular spelling to sound correspondence [54]. The same can be said for languages that incorporate phonograms or ideograms (e.g., Chinese, Japanese, Korean or Vietnamese) which could be more context-dependent or invoke greater imageability in reading. Results may therefore differ significantly in NART-like tests developed for more "transparent" languages. Fourth, item-level analyses were conducted on a subsample of participants for which item-level AmNART data (as opposed to only total AmNART score) was available (195 HC, 323 LMCI, 156 AD). Additionally, 20 words with missing values in age of acquisition, objective lexical frequency, concreteness, and/or phonological neighbourhood had to be excluded from item-level analyses. Fifth, the use of the BNT as unique semantic test has certain limits, as picture naming involves distinct cognitive processes that are not limited to semantics. Involved are visual analysis of the picture, recognition of the stimulus as familiar, activation of the semantic representation of the object via the semantic system, a lexical-semantic process which directs selection and retrieval of semantic information in a task appropriate way, modality-independent lexical access to the phonological word form of the object, that is to say the speech sounds used in the word; and the motor programming and articulation required for saying the word [134, 135]. This is important as reading models see the involvement of both lexical and semantic processing in correct reading of irregular words [77], these results should therefore not be interpreted as an involvement of semantics alone. Sixth, it is important to note that the AD dementia population recruited in the ADNI study is at relatively early stages of the disease (i.e. MMSE score between 20 and 26 and CDR score=0.5 or 1). Future studies could investigate irregular word reading in later stages of the disease.

Conclusions

Measuring cognitive decline can be particularly challenging for clinicians when considering that diseases as insidious as AD dementia may be involved. Cognitive decline will more often than not have to be estimated post-hoc,

Marier *et al. Alzheimer's Research & Therapy*   (2024) 16:96

Page 15 of 19

blind to an individual's objective baseline performances. The first assessment, where only one time point is available, could prove critical to any intervention against the disease and its progression. Currently, clinicians have to rely on subjective complaints, demographically-adjusted norms of cognitive performance and repeated measures. The results of this study lend support to the idea that irregular word reading tests do not provide an accurate estimate of premorbid IQ in the MCI-AD populations as it appears irregular word reading performances significantly declines in this population and are related to semantic impairments correlated to hippocampal and ATL volume loss. Relying on these estimates could lead clinicians to underestimate cognitive decline in people with those conditions. Premorbid estimates should rely on more crystalized forms of intelligence that are uncorrelated to disease severity as evidenced by longitudinal studies in clinically diverse populations.

Abbreviations

AD	Alzheimer's disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
AmNART	American National Adult Reading Test
ANOVA	Analysis of variance
ATL	Anterior temporal lobe
RAVLT	Rey Auditory Verbal Learning Test
BNT	Boston Naming Test
CDR	Clinical Dementia Rating
DBM	Deformation-based morphometry
ELP	English lexicon project
EMCI	Early mild cognitive impairment
FDR	False Discovery Rate
HC	Healthy controls
IQ	Intellectual quotient
LD	Levenshtein distance
LM II	Logical Memory II subscale
LMCI	Late mild cognitive impairment
MCI	Mild cognitive impairment
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
MRI	Magnetic resonance imaging
NART	National Adult Reading Test
OLD20	Orthographic Levenshtein distance to the 20 closest neighbors
PET	Positron emission tomography
PLD20	Phonologic Levenshtein distance to the 20 closest neighbors
ROI	Region of interest
SCD	Subjective cognitive decline
T1w	T1-weighted
WMS-R	Wechsler Memory Scale—Revised

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13195-024-01438-3.

Additional file 1: Supplementary Table 1. Contrasts between diagnostic categories on AmNART total error scores.

Authors' contributions

AM: Conceptualization (support); Data curation (lead); formal analysis (lead); methodology (support); software (lead); visualization (lead); writing – original draft preparation; writing – review & editing (support). MD: Formal analysis (support); data curation (support); formal analysis (support); visualization (support); writing – review & editing (support). FB: Conceptualization (support); data curation (support); formal analysis (support); methodology (support); resources (support); software (support); supervision (support); writing – review & editing (support). MM: Conceptualisation (lead); project administration (lead); methodology (lead); resources (lead); software (support); supervision (lead); writing – review & editing (lead).

Availability of data and materials

Data analysed in this study were acquired from the ADNI database (http://adni.loni.usc.edu). All ADNI data are shared without embargo through the Laboratory of Neuro Imaging and Data Archive. Interested scientists may obtain access to ADNI imaging, clinical, genomic, and biomarker data for the purposes of scientific investigation, teaching, or planning clinical research studies. Access is contingent on adherence to the ADNI Data Use Agreement and the publications' policies. For further information, please refer to the ADNI website (http://adni.loni.usc.edu/data-samples/access-data/).

Declarations

Ethics approval and consent to participate

The ADNI study follows the ethical standards of the institutional research committees and adheres to the principles outlined in the 1975 Helsinki Declaration and its later amendments. All study participants, authorized representatives, and study partners have provided written informed consent, and each participating site of ADNI has obtained the necessary ethical permits. More details can be found at adni.loni.usc.edu.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Douglas Research Centre & Department of Psychiatry, McGill University, 6875 Boulevard LaSalle, Montréal, QC H4H 1R3, Canada. [2]Department of Psychology, Université de Montréal, Succursale Centre-Ville, Montréal, QC C.P. 6128, H3C 3J7, Canada. [3]LPC, Aix Marseille University, CNRS, Marseille, France.

Marier *et al. Alzheimer's Research & Therapy*   (2024) 16:96

Page 16 of 19

References

1. Alzheimer's disease facts and figures. (2023). *Alzheimer's & Dementia, 19*(4), 1598–1695. https://doi.org/10.1002/alz.13016.
2. Jessen F, Amariglio RE, van Boxtel M, Breteler M, Ceccaldi M, Chételat G, Dubois B, Dufouil C, Ellis KA, van der Flier WM, Glodzik L, van Harten AC, de Leon MJ, McHugh P, Mielke MM, Molinuevo JL, Mosconi L, Osorio RS, Perrotin A, Wagner M. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. Alzheimers & Dementia. 2014;10(6):844–52. https://doi.org/10.1016/j.jalz.2014.01.001.
3. Mitchell AJ, Beaumont H, Ferguson D, Yadegarfar M, Stubbs B. Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: meta-analysis. Acta Psychiatr Scand. 2014;130(6):439–51. https://doi.org/10.1111/acps.12336.
4. Visser PJ, Verhey F, Knol DL, Scheltens P, Wahlund LO, Freund-Levi Y, Tsolaki M, Minthon L, Wallin AK, Hampel H, Bürger K, Pirttila T, Soininen H, Rikkert MO, Verbeek MM, Spiru L, Blennow K. Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: a prospective cohort study. The Lancet Neurology. 2009;8(7):619–27. https://doi.org/10.1016/S1474-4422(09)70139-5.
5. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dementia. 2011;7(3):270–9. https://doi.org/10.1016/j.jalz.2011.03.008.
6. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011;7(3):263–9. https://doi.org/10.1016/j.jalz.2011.03.005.
7. Beardsall L, Huppert FA. Improvement in NART word reading in demented and normal older persons using the Cambridge Contextual Reading Test. J Clin Exp Neuropsychol. 1994;16(2):232–42. https://doi.org/10.1080/01688639408402634.
8. Blair JR, Spreen O. Predicting premorbid IQ: A revision of the national adult reading test. Clinical Neuropsychologist. 1989;3(2):129–36. https://doi.org/10.1080/13854048908403285.
9. Del Ser T, González-Montalvo JI, Martínez-EspinosaDelgado-VillapalosBermejo SCF. Estimation of Premorbid Intelligence in Spanish People with the Word Accentuation Test and Its Application to the Diagnosis of Dementia. Brain Cogn. 1997;33(3):343–56. https://doi.org/10.1006/brcg.1997.0877.
10. Grober E, Sliwinski M. Development and validation of a model for estimating premorbid verbal intelligence in the elderly. J Clin Exp Neuropsychol. 1991;13(6):933–49. https://doi.org/10.1080/01688639108405109.
11. Nelson, H. E. (1982). *Nelson Adult Reading Test manual*. The National Hospital for Nervous Disease.
12. Nelson, H. E. (1982). *National Adult Reading Test (NART): For the assessment of premorbid intelligence in patients with dementia: Test manual*. Windsor: NFER-Nelson.
13. Schmand B, Bakker D, Saan R, Louman J. De Nederlandse Leestest voor Volwassenen: een maat voor het premorbide intelligentieniveau [The Dutch Reading Test for Adults: a measure of premorbid intelligence level]. Tijdschr Gerontol Geriatr. 1991;22(1):15–9.
14. Vaskinn A, Sundet K. Estimering av premorbid IQ: En norsk versjon av National Adult Reading Test/Estimating premorbid IQ: A Norwegian version of National Adult Reading Test. Tidsskrift for Norsk Psykologforening. 2001;38(12):1133–40.
15. Alves L, Simões MR, Martins C. The estimation of premorbid intelligence levels among Portuguese speakers: the Irregular Word Reading Test (TeLPI). Arch Clin Neuropsychol. 2012;27(1):58–68. https://doi.org/10.1093/arclin/acr103.
16. Ginsberg JP, Risser AH, Purisch AD, Hamilton W, Burns TG. BOOK AND TEST REVIEWS Applied Neuropsychology. 2003;10(3):182–90. https://doi.org/10.1207/S15324826AN1003_08.
17. Mackinnon A, Mulligan R. Estimation de l'intelligence prémorbide chez les francophones [The estimation of premorbid intelligence levels in French speakers]. L'Encephale. 2005;31(1 Pt 1):31–43. https://doi.org/10.1016/s0013-7006(05)82370-x.
18. Matsuoka K, Uno M, Kasai K, Koyama K, Kim Y. Estimation of premorbid IQ in individuals with Alzheimer's disease using Japanese ideographic script (Kanji) compound words: Japanese version of National Adult Reading Test. Psychiatry and Clinical Neuroscience. 2006;60(3):332–9. https://doi.org/10.1111/j.1440-1819.2006.01510.x.
19. Rolstad S, Nordlund A, Gustavsson MH, Eckerström C, Klang O, Hansen S, Wallin A. The Swedish National Adult Reading Test (NART-SWE): a test of premorbid IQ. Scand J Psychol. 2008;49(6):577–82. https://doi.org/10.1111/j.1467-9450.2008.00677.x.
20. Starkey NJ, Halliday T. Development of the New Zealand Adult Reading Test (NZART): Preliminary findings. N Z J Psychol. 2011;40(3):129–41.
21. Tallberg I-M, Wenneborg K, Almkvist O. Reading words with irregular decoding rules: A test of premorbid cognitive function? Scand J Psychol. 2006;47(6):531–9. https://doi.org/10.1111/j.1467-9450.2006.00547.x.
22. Yi D, Seo EH, Han JY, Sohn BK, Byun MS, Lee JH, Choe YM, Ahn S, Woo JI, Jun J, Lee DY. Development of the Korean Adult Reading Test (KART) to estimate premorbid intelligence in dementia patients. PLoS ONE. 2017;12(7): e0181523. https://doi.org/10.1371/journal.pone.0181523.
23. Nelson HE, McKenna P. The Use of Current Reading Ability in the Assessment of Dementia. British Journal of Social and Clinical Psychology. 1975;14(3):259–67. https://doi.org/10.1111/j.2044-8260.1975.tb00178.x.
24. Horn JL, Cattell RB. Refinement and test of the theory of fluid and crystallized general intelligences. J Educ Psychol. 1966;57(5):253–70. https://doi.org/10.1037/h0023816.
25. Sternberg RJ, Kaufman SB. The Cambridge Handbook of Intelligence. Cambridge University Press. 2011. https://doi.org/10.1017/CBO9780511977244.
26. Aichele S, Rabbitt P, Ghisletta P. Life span decrements in fluid intelligence and processing speed predict mortality risk. Psychol Aging. 2015;30(3):598–612. https://doi.org/10.1037/pag0000035.
27. McArdle JJ, Ferrer-Caja E, Hamagami F, Woodcock RW. Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. Dev Psychol. 2002;38(1):115–42. https://doi.org/10.1037/0012-1649.38.1.115.
28. Nelson HE, O'Connell A. Dementia: the estimation of premorbid intelligence levels using the New Adult Reading Test. Cortex. 1978;14(2):234–44. https://doi.org/10.1016/s0010-9452(78)80049-5.
29. Bright P, Hale E, Gooch VJ, Myhill T, van der Linde I. The National Adult Reading Test: restandardisation against the Wechsler Adult Intelligence Scale—Fourth edition. Neuropsychol Rehabil. 2018;28(6):1019–27. https://doi.org/10.1080/09602011.2016.1231121.
30. Uttl B. North American Adult Reading Test: Age Norms, Reliability, and Validity. J Clin Exp Neuropsychol. 2002;24(8):1123–37. https://doi.org/10.1076/jcen.24.8.1123.8375.
31. Crawford JR, Parker DM, Besson JAO. Estimation of Premorbid Intelligence in Organic Conditions. Br J Psychiatry. 1988;153(2):178–81. https://doi.org/10.1192/bjp.153.2.178.
32. Cummings JL, Houlihan JP, Hill MA. The pattern of reading deterioration in dementia of the Alzheimer type: Observations and implications. Brain Lang. 1986;29(2):315–23. https://doi.org/10.1016/0093-934X(86)90051-9.
33. Hart S, Smith CM, Swash M. Assessing intellectual deterioration. Br J Clin Psychol. 1986;25(Pt 2):119–24. https://doi.org/10.1111/j.2044-8260.1986.tb00680.x.
34. Johnstone B, Callahan CD, Kapila CJ, Bouman DE. The comparability of the WRAT-R reading test and NAART as estimates of premorbid intelligence in neurologically impaired patients. Arch Clin Neuropsychol. 1996;11(6):513–9. https://doi.org/10.1016/0887-6177(96)82330-4.
35. Maddrey AM, Cullum CM, Weiner MF, Filley CM. Premorbid intelligence estimation and level of dementia in Alzheimer's disease. J Int Neuropsychol Soc. 1996;2(6):551–5. https://doi.org/10.1017/S1355617700001727.

36. Nebes RD, Martin DC, Horn LC. Sparing of semantic memory in Alzheimer's disease. J Abnorm Psychol. 1984;93(3):321–30. https://doi.org/10.1037//0021-843x.93.3.321.

37. O'Carroll RE, Gilleard CJ. Estimation of premorbid intelligence in dementia. Br J Clin Psychol. 1986;25(2):157–8. https://doi.org/10.1111/j.2044-8260.1986.tb00690.x.

38. Raymer AM, Berndt RS. Reading lexically without semantics: Evidence from patients with probable Alzheimer's disease. J Int Neuropsychol Soc. 1996;2(4):340–9. https://doi.org/10.1017/S1355617700001363.

39. Schlosser D, Ivison D. Assessing memory deterioration with the wechsler memory scale, the national adult Reading Test, and the Schonell Graded Word Reading Test. J Clin Exp Neuropsychol. 1989;11(6):785–92. https://doi.org/10.1080/01688638908400935.

40. Sharpe K, O'Carroll R. Estimating premorbid intellectual level in dementia using the National Adult Reading Test: A Canadian study. Br J Clin Psychol. 1991;30(4):381–4. https://doi.org/10.1111/j.2044-8260.1991.tb00962.x.

41. Smith GE, Bohac DL, Ivnik RJ, Malec JF. Using word recognition tests to estimate premorbid IQ in early dementia: Longitudinal data. J Int Neuropsychol Soc. 1997;3(6):528–33. https://doi.org/10.1017/S1355617797005286.

42. Alves L, Simões MR, Martins C, Freitas S, Santana I. TeLPI performance in subjects with mild cognitive impairment and Alzheimer disease: a validation study. Alzheimer Dis Assoc Disord. 2013;27(4):324–9. https://doi.org/10.1097/WAD.0b013e31827bdc8c.

43. Bright P, Jaldow ELI, Kopelman MD. The National Adult Reading Test as a measure of premorbid intelligence: A comparison with estimates derived from demographic variables. J Int Neuropsychol Soc. 2002;8(6):847–54. https://doi.org/10.1017/S1355617702860131.

44. Law R, O'Carroll RE. A comparison of three measures of estimating premorbid intellectual level in dementia of the Alzheimer type. Int J Geriatr Psychiatry. 1998;13(10):727–30. https://doi.org/10.1002/(sici)1099-1166(1998100)13:10%3c727::aid-gps851%3e3.0.co;2-2.

45. Luzzatti C, Laiacona M, Agazzi D. Multiple patterns of writing disorders in dementia of the Alzheimer type and their evolution. Neuropsychologia. 2003;41(7):759–72. https://doi.org/10.1016/s0028-3932(02)00328-7.

46. McGurn B, Starr JM, Topfer JA, Pattie A, Whiteman MC, Lemmon HA, Whalley LJ, Deary IJ. Pronunciation of irregular words is preserved in dementia, validating premorbid IQ estimation. Neurology. 2004;62(7):1184–6. https://doi.org/10.1212/01.wnl.0000103169.80910.8b.

47. Paolo AM, Tröster AI, Ryan JJ, Koller WC. Comparison of NART and Barona demographic equation premorbid IQ estimates in Alzheimer's disease. J Clin Psychol. 1997;53(7):713–22. https://doi.org/10.1002/(SICI)1097-4679(199711)53:7%3c713::AID-JCLP8%3e3.0.CO;2-I.

48. Hughes JC, Graham N, Patterson K, Hodges JR. Dysgraphia in mild dementia of Alzheimer's type. Neuropsychologia. 1997;35(4):533–45. https://doi.org/10.1016/s0028-3932(96)00102-9.

49. O'Carroll RE, Prentice N, Murray C, Van Beck M, Ebmeier KP, Goodwin GM. Further Evidence that Reading Ability is not Preserved in Alzheimer's Disease. Br J Psychiatry. 1995;167(5):659–62. https://doi.org/10.1192/bjp.167.5.659.

50. Patterson KE, Graham N, Hodges JR. Reading in dementia of the Alzheimer type: A preserved ability? Neuropsychology. 1994;8(3):395–407. https://doi.org/10.1037/0894-4105.8.3.395.

51. Rapcsak SZ, Arthur SA, Bliklen DA, Rubens AB. Lexical Agraphia in Alzheimer's Disease. Arch Neurol. 1989;46(1):65–8. https://doi.org/10.1001/archneur.1989.00520370067021.

52. Storandt M, Stone K, LaBarge E. Deficits in reading performance in very mild dementia of the Alzheimer type. Neuropsychology. 1995;9(2):174–6. https://doi.org/10.1037/0894-4105.9.2.174.

53. Conway SC, O'Carroll RE. An evaluation of the Cambridge Contextual Reading Test (CCRT) in Alzheimer's disease. Br J Clin Psychol. 1997;36(4):623–5. https://doi.org/10.1111/j.2044-8260.1997.tb01267.x.

54. Colombo L, Brivio C, Benaglio I, Siri S, Cappa SF. Alzheimer Patients' Ability to Read Words with Irregular Stress. Cortex. 2000;36(5):703–14. https://doi.org/10.1016/S0010-9452(08)70547-1.

55. Colombo L, Fonti C, Cappa S. The impact of lexical-semantic impairment and of executive dysfunction on the word reading performance of patients with probable Alzheimer dementia. Neuropsychologia. 2004;42(9):1192–202. https://doi.org/10.1016/j.neuropsychologia.2004.02.004.

56. Glosser G, Grugan P, Friedman RB. Comparison of reading and spelling in patients with probable Alzheimer's disease. Neuropsychology. 1999;13(3):350–8. https://doi.org/10.1037/0894-4105.13.3.350.

57. McFarlane J, Welch J, Rodgers J. Severity of Alzheimer's disease and effect on premorbid measures of intelligence. Br J Clin Psychol. 2006;45(Pt 4):453–63. https://doi.org/10.1348/014466505x71245.

58. Noble K, Glosser G, Grossman M. Oral Reading in Dementia. Brain Lang. 2000;74(1):48–69. https://doi.org/10.1006/brln.2000.2330.

59. Pestell S, Shanks MF, Warrington J, Venneri A. Quality of Spelling Breakdown in Alzheimer's Disease is Independent of Disease Progression. J Clin Exp Neuropsychol. 2000;22(5):599–612. https://doi.org/10.1076/1380-3395(200010)22:5;1-9;FT599.

60. Taylor R. National Adult Reading Test performance in established dementia. Arch Gerontol Geriatr. 2000;29(3):291–6. https://doi.org/10.1016/S0167-4943(99)00042-4.

61. Weekes B. Oral reading in Chinese: evidence from dementia of the Alzheimer's type. Int J Lang Commun Disord. 2000;35(4):543–59. https://doi.org/10.1080/136828200750001287.

62. Cockburn J, Keene J, Hope T, Smith P. Progressive decline in NART score with increasing dementia severity. J Clin Exp Neuropsychol. 2000;22(4):508–17. https://doi.org/10.1076/1380-3395(200008)22:4;1-0;ft508.

63. Fromm D, Holland AL, Nebes RD, Oakley MA. A longitudinal study of word-reading ability in Alzheimer's disease: evidence from the National Adult Reading Test. Cortex. 1991;27(3):367–76. https://doi.org/10.1016/s0010-9452(13)80032-9.

64. Grober E, Hall CB, Lipton RB, Zonderman AB, Resnick SM, Kawas C. Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. J Int Neuropsychol Soc. 2008;14(2):266–78. https://doi.org/10.1017/S1355617708080302.

65. Paque L, Warrington EK. A longitudinal study of reading ability in patients suffering from dementia. J Int Neuropsychol Soc. 1995;1(6):517–24. https://doi.org/10.1017/S1355617700000643.

66. Pavlik VN, Doody RS, Massman PJ, Chan W. Influence of Premorbid IQ and Education on Progression of Alzheimer's Disease. Dement Geriatr Cogn Disord. 2006;22(4):367–77. https://doi.org/10.1159/000095640.

67. Strain E, Patterson K, Graham N, Hodges JR. Word reading in Alzheimer's disease: cross-sectional and longitudinal analyses of response time and accuracy data. Neuropsychologia. 1998;36(2):155–71. https://doi.org/10.1016/S0028-3932(97)00092-4.

68. Taylor KI, Salmon DP, Rice VA, Bondi MW, Hill LR, Ernesto CR, Butters N. Longitudinal examination of american national adult reading test (AMNART) performance in dementia of the Alzheimer type (DAT): Validation and correction based on degree of cognitive decline. J Clin Exp Neuropsychol. 1996;18(6):883–91. https://doi.org/10.1080/01688639608408309.

69. Lowe DA, Rogers SA. Estimating Premorbid Intelligence among Older Adults: The Utility of the AMNART. J Aging Res. 2011;2011:428132–428132. https://doi.org/10.4061/2011/428132.

70. Weinborn M, Bucks RS, Sohrabi HR, Rainey-Smith SR, Brown BM, Gardener SL, Gozt A, Christensen D, Savage G, Laws SM, Taddei K, Maruff P, Robertson JS, Ellis KA, Ames D, Masters CL, Rowe CC, Martins RN. Trajectories of irregular word reading ability as a proxy for premorbid intelligence in Alzheimer's disease, mild cognitive impairment, and healthy aging: A longitudinal study. Psychol Assess. 2018;30(10):1308–16. https://doi.org/10.1037/pas0000565.

71. Brambati SM, Ogar J, Neuhaus J, Miller BL, Gorno-Tempini ML. Reading disorders in primary progressive aphasia: A behavioral and neuroimaging study. Neuropsychologia. 2009;47(8):1893–900. https://doi.org/10.1016/j.neuropsychologia.2009.02.033.

72. Seidenberg MS, McClelland JL. A distributed, developmental model of word recognition and naming. Psychol Rev. 1989;96(4):523–68. https://doi.org/10.1037/0033-295X.96.4.523.

73. Patterson K, Hodges JR. Deterioration of word meaning: Implications for reading. Neuropsychologia. 1992;30(12):1025–40. https://doi.org/10.1016/0028-3932(92)90096-5.

74. Chapleau M, Wilson MA, Potvin K, Harvey-Langton A, Montembeault M, Brambati SM. Word reading aloud skills: their positive redefinition

through ageing. J Res Reading. 2017;40:297–312. https://doi.org/10.1111/1467-9817.12065.

75. Coltheart M. Acquired dyslexias and the computational modelling of reading. Cogn Neuropsychol. 2006;23(1):96–109. https://doi.org/10.1080/02643290500202649.

76. Plaut DC, McClelland JL, Seidenberg MS, Patterson K. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. Psychol Rev. 1996;103(1):56–115. https://doi.org/10.1037/0033-295X.103.1.56.

77. Taylor JSH, Rastle K, Davis MH. Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. Psychol Bull. 2013;139(4):766–91. https://doi.org/10.1037/a0030266.

78. Taylor JSH, Duff FJ, Woollams AM, Monaghan P, Ricketts J. How Word Meaning Influences Word Reading. Curr Dir Psychol Sci. 2015;24(4):322–8. https://doi.org/10.1177/0963721415574980.

79. Woollams AM, Ralph MA, Plaut DC, Patterson K. SD-squared: on the association between semantic dementia and surface dyslexia. Psychol Rev. 2007;114(2):316–39. https://doi.org/10.1037/0033-295x.114.2.316.

80. Predovan D, Gandini D, Montembeault M, Rouleau I, Bherer L, Joubert S, Brambati SM. Loss of person-specific knowledge in Alzheimer's disease: evidence from priming. Neurocase. 2014;20(3):263–8. https://doi.org/10.1080/13554794.2013.770876.

81. Joubert S, Vallet GT, Montembeault M, Boukadi M, Wilson MA, Laforce RJ, Rouleau I, Brambati SM. Comprehension of concrete and abstract words in semantic variant primary progressive aphasia and Alzheimer's disease: A behavioral and neuroimaging study. Brain Lang. 2017;170:93–102. https://doi.org/10.1016/j.bandl.2017.04.004.

82. Joubert S, Gardy L, Didic M, Rouleau I, Barbeau EJ. A Meta-Analysis of Semantic Memory in Mild Cognitive Impairment. Neuropsychol Rev. 2021;31(2):221–32. https://doi.org/10.1007/s11065-020-09453-5.

83. Montembeault M, Chapleau M, Jarret J, Boukadi M, Laforce R Jr, Wilson MA, Rouleau I, Brambati SM. Differential language network functional connectivity alterations in Alzheimer's disease and the semantic variant of primary progressive aphasia. Cortex; a journal devoted to the study of the nervous system and behavior. 2019;117:284–98. https://doi.org/10.1016/j.cortex.2019.03.018.

84. Montembeault M, Brambati SM, Joubert S, Boukadi M, Chapleau M, Laforce RJ, Wilson MA, Macoir J, Rouleau I. Naming unique entities in the semantic variant of primary progressive aphasia and Alzheimer's disease: Towards a better understanding of the semantic impairment. Neuropsychologia. 2017;95:11–20. https://doi.org/10.1016/j.neuropsychologia.2016.12.009.

85. Shimotake A, Matsumoto R, Ueno T, Kunieda T, Saito S, Hoffman P, Kikuchi T, Fukuyama H, Miyamoto S, Takahashi R, Ikeda A, Lambon Ralph MA. Direct Exploration of the Role of the Ventral Anterior Temporal Lobe in Semantic Memory: Cortical Stimulation and Local Field Potential Evidence From Subdural Grid Electrodes. Cereb Cortex. 2014;25(10):3802–17. https://doi.org/10.1093/cercor/bhu262.

86. Hoffman P, Lambon Ralph MA, Woollams AM. Triangulation of the neurocomputational architecture underpinning reading aloud. Proceedings of the National Academy of Sciences. 2015;112(28):E3719–28. https://doi.org/10.1073/pnas.1502032112.

87. Joyal M, Brambati SM, Laforce RJ, Montembeault M, Boukadi M, Rouleau I, Macoir J, Joubert S, Fecteau S, Wilson MA. The Role of the Left Anterior Temporal Lobe for Unpredictable and Complex Mappings in Word Reading. Front Psychol. 2017;8:517. https://doi.org/10.3389/fpsyg.2017.00517.

88. Ueno T, Meteyard L, Hoffman P, Murayama K. The Ventral Anterior Temporal Lobe has a Necessary Role in Exception Word Reading. Cereb Cortex. 2018;28(8):3035–45. https://doi.org/10.1093/cercor/bhy131.

89. Wilson MA, Joubert S, Ferré P, Belleville S, Ansaldo AI, Joanette Y, Rouleau I, Brambati SM. The role of the left anterior temporal lobe in exception word reading: Reconciling patient and neuroimaging findings. Neuroimage. 2012;60(4):2000–7. https://doi.org/10.1016/j.neuroimage.2012.02.009.

90. Chapleau M, Aldebert J, Montembeault M, Brambati SM. Atrophy in Alzheimer's Disease and Semantic Dementia: An ALE Meta-Analysis of Voxel-Based Morphometry Studies. Journal of Alzheimer's disease: JAD. 2016;54(3):941–55. https://doi.org/10.3233/JAD-160382.

91. Frings L, Klöppel S, Teipel S, Peters O, Frölich L, Pantel J, Schröder J, Gertz HJ, Arlt S, Heuser I, Kornhuber J, Wiltfang J, Maier W, Jessen F, Hampel H, Hüll M. Left anterior temporal lobe sustains naming in Alzheimer's dementia and mild cognitive impairment. Curr Alzheimer Res. 2011;8(8):893–901. https://doi.org/10.2174/156720511798192673.

92. Joubert S, Brambati SM, Ansado J, Barbeau EJ, Felician O, Didic M, Lacombe J, Goldstein R, Chayer C, Kergoat MJ. The cognitive and neural expression of semantic memory impairment in mild cognitive impairment and early Alzheimer's disease. Neuropsychologia. 2010;48(4):978–88. https://doi.org/10.1016/j.neuropsychologia.2009.11.019.

93. Hansson O, Seibyl J, Stomrud E, Zetterberg H, Trojanowski JQ, Bittner T, Lifke V, Corradini V, Eichenlaub U, Batrla R, Buck K, Zink K, Rabe C, Blennow K, Shaw LM, BioFINDER S, study group, & Alzheimer's Disease Neuroimaging Initiative,. CSF biomarkers of Alzheimer's disease concord with amyloid-β PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts. Alzheimer's Dementia. 2018;14(11):1470–81. https://doi.org/10.1016/j.jalz.2018.01.010.

94. Blennow K, Shaw LM, Stomrud E, Mattsson N, Toledo JB, Buck K, Wahl S, Eichenlaub U, Lifke V, Simon M, Trojanowski JQ, Hansson O. Predicting clinical decline and conversion to Alzheimer's disease or dementia using novel Elecsys Aβ(1–42), pTau and tTau CSF immunoassays. Sci Rep. 2019;9(1):19024. https://doi.org/10.1038/s41598-019-54204-z.

95. Balota DA, Yap MJ, Hutchison KA, Cortese MJ, Kessler B, Loftis B, Neely JH, Nelson DL, Simpson GB, Treiman R. The English Lexicon Project. Behav Res Methods. 2007;39(3):445–59. https://doi.org/10.3758/BF03193014.

96. Miller GA. WordNet: a lexical database for English. Commun ACM. 1995;38(11):39–41. https://doi.org/10.1145/219717.219748.

97. Yarkoni T, Balota D, Yap M. Moving beyond Coltheart's N: a new measure of orthographic similarity. Psychon Bull Rev. 2008;15(5):971–9. https://doi.org/10.3758/pbr.15.5.971.

98. Kuperman V, Stadthagen-Gonzalez H, Brysbaert M. Age-of-acquisition ratings for 30,000 English words. Behav Res Methods. 2012;44(4):978–90. https://doi.org/10.3758/s13428-012-0210-4.

99. Elsherif MM, Preece E, Catling JC. Age-of-acquisition effects: A literature review. J Exp Psychol Learn Mem Cogn. 2023;49(5):812–47. https://doi.org/10.1037/xlm0001215.

100. Juhasz BJ. Age-of-Acquisition Effects in Word and Picture Identification. Psychol Bull. 2005;131(5):684–712. https://doi.org/10.1037/0033-2909.131.5.684.

101. Brysbaert M, Warriner AB, Kuperman V. Concreteness ratings for 40 thousand generally known English word lemmas. Behav Res Methods. 2014;46(3):904–11. https://doi.org/10.3758/s13428-013-0403-5.

102. Mirman D, Magnuson JS. Attractor dynamics and semantic neighborhood density: processing is slowed by near neighbors and speeded by distant neighbors. J Exp Psychol Learn Mem Cogn. 2008;34(1):65–79. https://doi.org/10.1037/0278-7393.34.1.65.

103. Shaoul C, Westbury C. Exploring lexical co-occurrence space using HiDEx. Behavioral Research Methods. 2010;42(2):393–413. https://doi.org/10.3758/brm.42.2.393.

104. Folstein MF, Folstein SE, McHugh PR. Mini-mental state A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research. 1975;12(3):189–98. https://doi.org/10.1016/0022-3956(75)90026-6.

105. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc. 2005;53(4):695–9. https://doi.org/10.1111/j.1532-5415.2005.53221.x.

106. Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test*. APA PsycTests.

107. Partington JE, Leiter RG. Partington's Pathways Test. Psychological Service Center Journal. 1949;1:11–20.

108. Rey, A. (1958). *L'examen clinique en psychologie*. [The clinical examination in psychology]. Presses Universitaries De France.

109. Coupe P, Yger P, Prima S, Hellier P, Kervrann C, Barillot C. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. IEEE Trans Med Imaging. 2008;27(4):425–41. https://doi.org/10.1109/TMI.2007.906087.

Marier *et al. Alzheimer's Research & Therapy*   (2024) 16:96

Page 19 of 19

110. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans Med Imaging. 1998;17(1):87–97. https://doi.org/10.1109/42.668698.

111. Dadar M, Fonov VS, Collins DL, Alzheimer's Disease Neuroimaging Initiative,. A comparison of publicly available linear MRI stereotaxic registration techniques. Neuroimage. 2018;174:191–200. https://doi.org/10.1016/j.neuroimage.2018.03.025.

112. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal. 2008;12(1):26–41. https://doi.org/10.1016/j.media.2007.06.004.

113. Dadar M, Camicioli R, Duchesne S, Collins DL, Initiative ADN. The temporal relationships between white matter hyperintensities, neurodegeneration, amyloid beta, and cognition. Alzheimer's & dementia (Amsterdam, Netherlands). 2020;12(1): e12091. https://doi.org/10.1002/dad2.12091.

114. Okada K. Is Omega Squared Less Biased? a Comparison of Three Major Effect Size Indices in One-Way Anova. Behaviormetrika. 2013;40(2):129–47. https://doi.org/10.2333/bhmk.40.129.

115. Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences*, 4th ed. Lawrence Erlbaum Associates Publishers.

116. Bates, D., Machler, M., Bolker, B. and Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67, 1–48. https://doi.org/10.18637/jss.v067.i01.

117. Baayen R. Analyzing Linguistic Data: A Practical Introduction to Statistics using R. Cambridge: Cambridge University Press; 2008. https://doi.org/10.1017/CBO9780511801686.

118. Borghesani V, Battistella G, Mandelli ML, Welch A, Weis E, Younes K, Neuhaus J, Grinberg LT, Seeley WM, Spina S, Miller B, Miller Z, Gorno-Tempini ML. Regional and hemispheric susceptibility of the temporal lobe to FTLD-TDP type C pathology. NeuroImage Clinical. 2020;28: 102369. https://doi.org/10.1016/j.nicl.2020.102369.

119. Geraudie A, Pressman PS, Pariente J, Millanski C, Palser ER, Ratnasiri BM, Battistella G, Mandelli ML, Miller ZA, Miller BL, Sturm V, Rankin KP, Gorno-Tempini ML, Montembeault M. Expressive Prosody in Patients With Focal Anterior Temporal Neurodegeneration. Neurology. 2023. Advance online publication. doi: https://doi.org/10.1212/WNL.0000000000207516.

120. R Core Team (2023). *R: A language and environment for statistical computing.* (version 4.2.1) [software] R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

121. Deary IJ, Whalley LJ, Lemmon H, Crawford JR, Starr JM. The stability of individual differences in mental ability from childhood to old age: Follow-up of the 1932 Scottish Mental Survey. Intelligence. 2000;28(1):49–55. https://doi.org/10.1016/S0160-2896(99)00031-8.

122. Gustavson DE, Elman JA, Panizzon MS, Franz CE, Zuber J, Sanderson-Cimino M, Reynolds CA, Jacobson KC, Xian H, Jak AJ, Toomey R, Lyons MJ, Kremen WS. Association of baseline semantic fluency and progression to mild cognitive impairment in middle-aged men. Neurology. 2020;95(8):e973–83. https://doi.org/10.1212/WNL.0000000000010130.

123. Tchakoute CT, Sainani KL, Henderson VW, Investigators, f. t. R. i. A. s. D. Semantic Memory in the Clinical Progression of Alzheimer Disease. Cogn Behav Neurol. 2017;30(3):81–9. https://doi.org/10.1097/wnn.0000000000000131.

124. Rao YL, Ganaraja B, Murlimanju BV, Joy T, Krishnamurthy A, Agrawal A. Hippocampus and its involvement in Alzheimer's disease: a review. 3 Biotech. 2022;12(2):55. https://doi.org/10.1007/s13205-022-03123-4.

125. Chu LW. Alzheimer's disease: early diagnosis and treatment. Hong Kong Med J Xianggang yi xue za zhi. 2012;18(3):228–37.

126. Binder JR, Desai RH, Graves WW, Conant LL. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. Cereb Cortex. 2009;19(12):2767–96. https://doi.org/10.1093/cercor/bhp055.

127. Chapleau M, Montembeault M, Boukadi M, Bedetti C, Laforce R Jr, Wilson M, Brambati SM. The role of the hippocampus in the semantic variant of primary progressive aphasia: A resting-state fcMRI study. Hippocampus. 2019;29(11):1127–32. https://doi.org/10.1002/hipo.23156.

128. Duff MC, Brown-Schmidt S. The hippocampus and the flexible use and processing of language. Front Hum Neurosci. 2012;6:69. https://doi.org/10.3389/fnhum.2012.00069.

129. Piai V, Anderson KL, Lin JJ, Dewar C, Parvizi J, Dronkers NF, Knight RT. Direct brain recordings reveal hippocampal rhythm underpinnings of language processing. Proc Natl Acad Sci USA. 2016;113(40):11366–71. https://doi.org/10.1073/pnas.1603312113.

130. Geake JG, Hansen PC. Functional neural correlates of fluid and crystallized analogizing. Neuroimage. 2010;49(4):3489–97. https://doi.org/10.1016/j.neuroimage.2009.09.008.

131. Tadayon E, Pascual-Leone A, Santarnecchi E. Differential Contribution of Cortical Thickness, Surface Area, and Gyrification to Fluid and Crystallized Intelligence. Cereb Cortex. 2019;30(1):215–25. https://doi.org/10.1093/cercor/bhz082.

132. Birkenbihl C, Salimi Y, Domingo-Fernándéz D, Lovestoneconsortium SA, Fröhlich H, Hofmann-Apitius Initiative M. Evaluating the Alzheimer's disease data landscape. Alzheimers Dement (N Y). 2020;6(1):e12102. https://doi.org/10.1002/trc2.12102.

133. Tsoy E, Kiekhofer RE, Guterman EL, Tee BL, Windon CC, Dorsman KA, Lanata SC, Rabinovici GD, Miller BL, Kind AJH, Possin KL. Assessment of Racial/Ethnic Disparities in Timeliness and Comprehensiveness of Dementia Diagnosis in California. JAMA Neurol. 2021;78(6):657–65. https://doi.org/10.1001/jamaneurol.2021.0399.

134. DeLeon J, Gottesman RF, Kleinman JT, Newhart M, Davis C, Heidler-Gary J, Lee A, Hillis AE. Neural regions essential for distinct cognitive processes underlying picture naming. Brain. 2007;130(Pt 5):1408–22. https://doi.org/10.1093/brain/awm011.

135. Harry A, Crowe SF. Is the Boston Naming Test Still Fit For Purpose? Clin Neuropsychol. 2014;28(3):486–504. https://doi.org/10.1080/13854046.2014.892155.

Publisher's Note